



Réduction de la Complexité du Décodeur Sphérique grâce à l'Adaptation de la Décomposition QR

Sébastien Aubert, Fabienne Nouvel, Marc Soler

► To cite this version:

Sébastien Aubert, Fabienne Nouvel, Marc Soler. Réduction de la Complexité du Décodeur Sphérique grâce à l'Adaptation de la Décomposition QR. Conférence Nationale Manifestation des Jeunes Chercheurs Francophones dans le domaine des STIC MAJECSTIC 2008, Oct 2008, Marseille, France. 6 p. hal-00444356

HAL Id: hal-00444356

<https://hal.science/hal-00444356>

Submitted on 8 Jan 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Réduction de la Complexité du Décodeur Sphérique grâce à l'Adaptation de la Décomposition QR

Sébastien Aubert* — Fabienne Nouvel** — Marc Soler*

*Entreprise NXP founded by Philips
505, route des lucioles, CP 06560 Sophia-Antipolis

**Laboratoire IETR (INSA)
20, avenue des buttes de Coësmes, CP 35043 Rennes
sebastien.aubert@nxp.com

RÉSUMÉ. L'objet de cet article est de présenter une nouvelle technique de détection, le décodeur sphérique, qui utilise une Décomposition QR. Dans cet article, la Décomposition QR de Householder, qui permet d'obtenir des coefficients réels positifs sur la diagonale de \mathbf{R} , est étudiée, alors que l'approche usuelle emploie la Décomposition QR de Gram Schmidt. Même si cette modification provoque une augmentation de la complexité des calculs de la Décomposition QR par rapport à la technique existante, elle permet de réduire fortement la complexité du Décodeur Sphérique adapté, de telle sorte que la complexité globale est réduite.

L'objectif de l'étude est de quantifier la diminution de la complexité globale dans un contexte Long Term Evolution (LTE), le protocole utilisé pour la quatrième génération de téléphones portables, sur une cible Embedded Vector Processor (EVP).

ABSTRACT. In this paper, we propose a new detection technique, associated with a QR Decomposition-based Sphere Decoder. The exposed Householder-based QR Decomposition, with \mathbf{R} matrix real diagonal elements, differ from a previous Gram Schmidt-based QR Decomposition. Although this change implies an increase of the QR Decomposition complexity compared to the previous technique, it induces a decrease of the adapted Sphere Decoder complexity, and overall complexity decreases.

In this paper, we quantify the overall complexity decrease in Long Term Evolution (LTE) context, in particular with an Embedded Vector Processor (EVP) target.

MOTS-CLÉS : MIMO, Décodage Sphérique, Décomposition QR, Householder, EVP.

KEYWORDS: MIMO, Sphere Decoder, QR Decomposition, Householder, EVP.

1. INTRODUCTION

Dans le contexte des communications numériques mobiles, les systèmes multi antennes *Multiple Input Multiple Output* (MIMO) sont constitués de plusieurs antennes émettrices et plusieurs antennes réceptrices aux entrées et sorties d'un canal de transmission. Ceux-ci permettent d'obtenir des augmentations importantes de la capacité et de la qualité de la transmission en termes de Taux d'Erreur Binaire (TEB). Cependant, cela implique une augmentation importante de la complexité de traitement, notamment en réception. En particulier, la séparation des flots de données transmis réalisée au moment de l'égalisation, qui a pour but de compenser les effets du canal, tient une part importante dans cette augmentation. Il s'agit d'un sujet de recherche essentiel dans la téléphonie mobile, puisque dans tout système embarqué chaque économie sur le gain de complexité global permet une réduction de la consommation, c'est-à-dire une plus grande autonomie, ou une économie sur la qualité des composants due à une diminution des contraintes.

2. ÉTUDE PRÉALABLE

2.1. Description du système

Soit un système MIMO $N_T \times N_R$, constitué de N_T antennes émettrices et de N_R antennes réceptrices. Le flot de données est démultiplexé en $k = \min(N_T, N_R)$ sous-flots secondaires parallèles qui sont ensuite transmis.

Soit $\mathbf{x} \in \mathbb{C}^k$ le vecteur des symboles transmis à un instant donné, tel que $\mathbf{x} = (x_0 \dots x_{k-1})^T$, $x_i \in \mathcal{S}_{modulation}$ le dictionnaire symboles de la modulation. En considérant un canal à évanouissements plats dans les bandes de fréquences du symbole *Orthogonal Frequency Division Multiplexing* (OFDM) (?) utilisées et une synchronisation parfaite, le vecteur des symboles reçus, noté $\mathbf{y} \in \mathbb{C}^{N_R}$, s'exprime sous forme matricielle : $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}$, avec $\mathbf{H} \in \mathbb{C}^{N_R \times N_T}$ la matrice de canal supposée connue à la réception, de coefficients h_{ij} qui caractérisent les gains des sous-canaux de l'antenne émettrice j vers l'antenne réceptrice i , tels que $h_{ij} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \sigma_H^2)$ où $\mathcal{N}_{\mathbb{C}}$ est la loi normale complexe. $\mathbf{n} \in \mathbb{C}^{N_R}$ est le vecteur de bruit additif $\mathbf{n} = (n_0 \dots n_{N_R-1})^T$ tel que $n_i \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \sigma_N^2)$. La problématique est alors de retrouver le vecteur des symboles transmis le plus correctement possible. Dans la suite de l'article, on considèrera le cas d'un système MIMO 2×2 .

2.2. Technique du Maximum de Vraisemblance

Le canal introduit des déformations et des interférences entre antennes, représentées par les gains h_{ij} tels que $i \neq j$, qui font chuter les performances du système. L'étape d'égalisation a pour objectif de réduire ces effets voire de les compenser, tandis que l'étape de détection consiste à déterminer le symbole qui a été transmis avec la plus grande probabilité. Les deux techniques nécessitent la connaissance du

canal en réception.

Des techniques d'égalisation linéaire comme le *Zero Forcing* (ZF), qui minimise les interférences entre symboles, et le *Minimum Mean Square Error* (MMSE), qui minimise l'erreur quadratique moyenne entre le signal transmis et son estimation, peuvent être utilisées. Ces techniques présentent l'avantage d'une complexité faible, au détriment des performances. L'algorithme de recherche du Maximum de Vraisemblance (MV), exact et exhaustif, représente la solution optimale en termes de Taux d'Erreur par Symbole et permet donc d'obtenir les meilleures performances. Il est cependant associé à une complexité importante, exponentielle, ce qui est difficilement réalisable dans le contexte des systèmes embarqués. Ce dilemme justifie l'introduction du Décodeur Sphérique (DS) en tant que technique de recherche efficace en termes de complexité et permettant d'obtenir des performances proches de celles du MV pour les systèmes MIMO.

3. DESCRIPTION TECHNIQUE

3.1. Description d'un algorithme du Décodeur Sphérique

L'algorithme utilisé permet un compromis entre les deux aspects décrits précédemment. Ses performances en termes de TEB sont très proches voire identiques à celles obtenues avec le MV alors que sa complexité est comparativement plus limitée. On note que dans le contexte *Long Term Evolution* (LTE) (?), les données sont transmises dans des formats de modulation *Quaternary Phase Shift Keying* (QPSK) ou *Quadrature Amplitude Modulation* à 16 ou 64 états (16-QAM ou 64-QAM).

3.1.1. Rayon de l'hypersphère

Un système MIMO peut être considéré comme un espace de dimension le nombre d'antennes émettrices, noté n , et doté de la norme euclidienne. Le DS considère les symboles situés à l'intérieur d'une hypersphère (sphère de dimension $n - 1$) de rayon d^2 . On dit qu'ils vérifient la contrainte de la sphère (?).

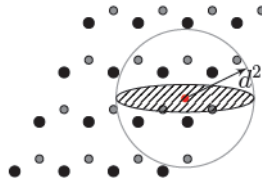


Figure 1. Contrainte de la sphère dans le cas d'une modulation 16-QAM sur deux antennes émettrices.

Une des principales difficultés de cette méthode est l'initialisation du rayon de cette sphère qui aura une très forte influence sur les résultats. Un rayon trop petit pour le système entraîne une chute des performances rédhibitoire, car trop peu de symboles possibles sont étudiés (voire aucun). Un rayon trop grand pour le système entraîne une trop grande complexité, proche de celle du MV et on perd l'intérêt de cette technique. Une possibilité est d'utiliser le rétrécissement du rayon d^2 et d'ordonner les symboles à étudier selon leurs *Partial Euclidian Distance* (PED) avec le symbole reçu croissantes, qui correspondent aux distances euclidiennes associées à un symbole du vecteur de symboles, afin de ne pas fixer de valeur initiale du rayon (?). Cette valeur est adaptative et on lui attribue la valeur $d^2 = \infty$ en début de recherche.

3.1.2. Algorithme de parcours en profondeur

La recherche du symbole optimal peut être représentée sous la forme d'un algorithme de parcours en profondeur d'un arbre (?). La procédure qui permet d'obtenir le symbole optimal consiste à démarrer la recherche sur les nœuds du niveau supérieur et vérifier si le premier nœud de niveau directement inférieur entraîne une distance euclidienne totale qui vérifie la contrainte de la sphère ou non. La recherche continue en examinant les nœuds suivants.

3.2. Étude de la Décomposition QR

En algèbre linéaire, la Décomposition QR est une factorisation d'une matrice $\mathbf{H} \in \mathbb{C}^{n \times m}$ en un produit \mathbf{QR} , où $\mathbf{Q} \in \mathbb{C}^{n \times k}$ est une matrice orthonormale et $\mathbf{R} \in \mathbb{C}^{k \times m}$ une matrice triangulaire supérieure, et on a (?) :

$$\|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 \leq d^2 \Rightarrow \|\mathbf{Q}^H \mathbf{y} - \mathbf{R}\mathbf{x}\|^2 \leq d'^2 \Rightarrow \|\mathbf{y}_{eq} - \mathbf{R}\mathbf{x}\|^2 \leq d'^2 \quad [1]$$

On note $(\cdot)^H$ la transposée hermitienne, \mathbf{y}_{eq} le vecteur des symboles reçus égalisé et d' le rayon de la sphère tel que $d' \neq d$.

3.2.1. Principe de la Décomposition QR de Gram Schmidt

L'algorithme de Gram Schmidt (GS) est une procédure d'orthonormalisation. Le but est de déterminer une matrice $\mathbf{Q} = (\mathbf{q}_1 | \mathbf{q}_2)$ de changement de base orthonormale vers une matrice triangulaire supérieure \mathbf{R} , à partir d'une base $\mathbf{H} = (\mathbf{h}_1 | \mathbf{h}_2)$ (?).

Avec cette technique, on détermine en premier la matrice \mathbf{Q} , et on en déduit la matrice \mathbf{R} à l'aide des matrices \mathbf{Q} et \mathbf{H} . Par conséquent, on obtient une matrice \mathbf{R} triangulaire supérieure $\begin{pmatrix} r_{11} & r_{12} \\ 0 & r_{22} \end{pmatrix}$, avec $r_{11} \in \mathbb{R}_+$, et $(r_{12}, r_{22}) \in \mathbb{C}^2$.

La technique détaillée ci-après permet d'obtenir des coefficients diagonaux réels positifs.

3.2.2. Principe de la Décomposition QR de Householder

La réduction de Householder (HH) est une procédure de triangularisation. Ainsi, le but est de déterminer une matrice triangulaire supérieure \mathbf{R} .

Avec cette technique, on détermine en premier la matrice \mathbf{R} , et on en déduit la matrice \mathbf{Q} à l'aide des matrices \mathbf{R} et \mathbf{H} . Par conséquent, on obtient une matrice \mathbf{R} triangulaire supérieure $\begin{pmatrix} r_{11} & r_{12} \\ 0 & r_{22} \end{pmatrix}$, avec $(r_{11}, r_{22}) \in \mathbb{R}_+^2$, et $r_{12} \in \mathbb{C}$. Bien que cet algorithme entraîne une augmentation de la complexité à chaque Décomposition QR par rapport à l'algorithme de GS, la complexité du s'en trouve largement diminuée.

Dans la suite de l'article, il est fait état de la diminution de la complexité globale.

3.2.3. Algorithme de la Décomposition QR de Householder

On note : $\mathbf{H} = (\mathbf{h}_1 | \mathbf{h}_2) = \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix}$, $(h_{11}, h_{12}, h_{21}, h_{22}) \in \mathbb{C}^4$.

On peut démontrer les résultats suivants. Le résultat de la réduction de HH est :

$$\mathbf{Q} = \begin{pmatrix} \frac{h_{11}}{\|\mathbf{h}_1\|} & \frac{h_{12} - k \cdot h_{11}}{\|\mathbf{h}_2'\|} \\ \frac{h_{21}}{\|\mathbf{h}_1\|} & \frac{h_{22} - k h_{21}}{\|\mathbf{h}_2'\|} \end{pmatrix}, \mathbf{R} = \begin{pmatrix} \|\mathbf{h}_1\| & k \|\mathbf{h}_1\| \\ 0 & \|\mathbf{h}_2'\| \end{pmatrix} \quad [2]$$

Avec : $\mathbf{h}_2' = \mathbf{h}_2 - k \mathbf{h}_1$, $k = \frac{\mathbf{h}_1^H \mathbf{h}_2}{\mathbf{h}_1^H \mathbf{h}_1} \in \mathbb{C}$

Alors que le résultat de l'orthonormalisation de GS est :

$$\mathbf{Q} = \begin{pmatrix} \frac{h_{11}}{\|\mathbf{h}_1\|} & \frac{h_{21}^*}{\|\mathbf{h}_1\|} \\ \frac{h_{21}}{\|\mathbf{h}_1\|} & -\frac{h_{11}^*}{\|\mathbf{h}_1\|} \end{pmatrix}, \mathbf{R} = \begin{pmatrix} \|\mathbf{h}_1\| & h_{11}^* h_{12} + h_{21}^* h_{22} \\ 0 & h_{12} h_{21} - h_{11} h_{22} \end{pmatrix} \quad [3]$$

Le tableau ?? précise le nombre d'opérations de multiplication réelle par symbole (*MUL*) (notons le nombre d'opérations d'addition réelle par symbole *ADD*) des Décompositions QR de GS et de HH qui sont effectuées.

La mise en œuvre de HH pour la Décomposition QR induit une augmentation de

| | GS | HH |
|-----|----|----|
| MUL | 64 | 90 |

Tableau 1. Bilans de complexités de la Décomposition QR.

MUL. Cependant, cette augmentation est compensée lors du déroulement du DS adapté, particulièrement lorsque le nombre de branches à explorer croît. Ce point est abordé dans la section suivante.

3.2.4. Impact sur le Décodeur Sphérique

La complexité du DS dépend du nombre de symboles respectant la contrainte de la sphère étudiés pour désigner le symbole optimal, c'est-à-dire du nombre de branches

de l'arbre parcourues (cf section 3.1.2).

La complexité minimale du DS correspond au cas où une seule branche est explorée, mais ce cas de figure correspond à une égalisation ZF. Le cas d'une complexité maximale dépend de l'ordre de la modulation employée puisqu'il modifie le nombre de feuilles de l'arbre. Ceci permet d'en déduire que le nombre L d'explorations en profondeur de l'arbre est borné de la façon suivante : $2 \leq L \leq \text{card}\{\mathcal{S}_{\text{modulation}}\}$.

Le tableau ?? précise MUL global (Décomposition QR de HH et DS adapté, Décomposition QR de GS et DS classique) pour les deux techniques associées à différents nombres L d'explorations en profondeur de toute une branche de l'arbre. Les bilans

| L | GS | HH |
|-----|-----|-----|
| 2 | 132 | 144 |
| 5 | 216 | 216 |
| 10 | 396 | 336 |

Tableau 2. Bilans de complexités globales.

des complexités globales permettent de noter l'intérêt de HH associé au DS adapté, par rapport à GS associé au DS classique, en termes de complexité lorsque le nombre de branches explorées augmente.

3.3. Résultats

Il est possible de déterminer un seuil du nombre de branches explorées à partir duquel la complexité de la technique de HH associée au DS adapté devient inférieure à celle de GS associée au DS classique. La figure ?? permet d'observer que la com-

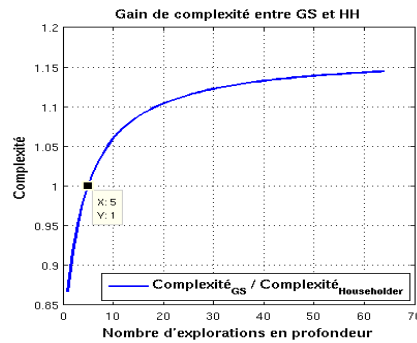


Figure 2. Gain de complexité entre les versions DS avec GS classique et DS adapté avec HH.

plexité globale du DS adapté associé à HH est inférieure ou égale à celle du DS classique associé à GS à partir de cinq branches explorées. Cette nouvelle technique offre

donc de meilleures performances dès que l'on utilise une constellation constituée de plus de cinq symboles, c'est-à-dire dès la modulation 16-QAM.

4. ASPECTS D'IMPLEMENTATION

4.1. *Le processeur Embedded Vector Processor (EVP)*

L'EVP est un processeur de traitement numérique du signal vectoriel (relatif aux données) capable de manipuler des vecteurs de données de 256 bits de large en un cycle d'horloge. Il s'agit d'un composant qui met en place le principe des Instructions à Données Multiples, associé à une architecture de bus de Harvard (deux bus distincts, pour les données et pour les instructions).

L'EVP est issu de l'entité recherche de *NXP founded by Philips*. Son architecture est décrite dans le détail (?) et est particulièrement adaptée aux algorithmes à parallélisme de données massif.

On peut montrer que, sur un EVP, le temps d'exécution d'une opération d'addition réelle est négligeable par rapport au temps d'exécution d'une opération de multiplication réelle à cause de son architecture superscalaire, car une addition peut presque toujours être réalisée en parallèle d'une multiplication (à condition que *MUL* soit supérieur ou égal à *ADD*). *MUL* nous permet ainsi de prédire les performances sur l'EVP en termes de Millions d'instructions par secondes (*Mips*).

4.2. *Quantification du gain de temps de traitement dans le contexte LTE*

Il est possible de présenter un exemple dans le contexte LTE (?) dans le cas de dix branches explorées, d'un symbole OFDM (?) qui occupe entre 72 et 1320 sous-porteuses, d'une durée de sous-fenêtre de 1 ms et de données transmises sur une seule antenne.

Afin de considérer les données seulement, il s'agit d'ignorer au moins un symbole pour le canal *Dedicated Physical Control Channel* ainsi qu'au moins douze symboles pilotes, appelés *Reference Symbols* à l'intérieur d'un *Radio Bloc*. Ceci implique, dans le contexte LTE où le nombre de sous-porteuses occupées est inférieur à 1320, que cette borne correspond en réalité à $1320 - 2 \times 12^2 = 1032$ (?).

En se plaçant dans le cas $L = 10$, la fréquence équivalente de fonctionnement nécessaire pour traiter les données d'une trame LTE passe de 357 MHz (*GS*) à 303 MHz (*HH*), correspondant respectivement à 396 *MUL* et 336 *MUL* (les *ADD* sont négligées). Les résultats de ce calcul permettent de réduire la fréquence du processeur de 54 MHz avec la nouvelle approche proposée.

4.3. Impact sur les performances du DS

Afin de vérifier les performances du DS obtenues dans les cas d'un prétraitement de la matrice \mathbf{H} à l'aide de GS d'une part et de HH d'autre part, on les compare avec celles du MV. Pour cela, on a mesuré leurs TEB en moyennant les taux d'erreur sur 100.000 itérations pour chaque type de modulation, au travers d'un canal MIMO 2×2 de Rayleigh, sur une seule porteuse, sans aucun codage et en utilisant des constellations du standard LTE. Les graphes montrent que toutes les courbes sont

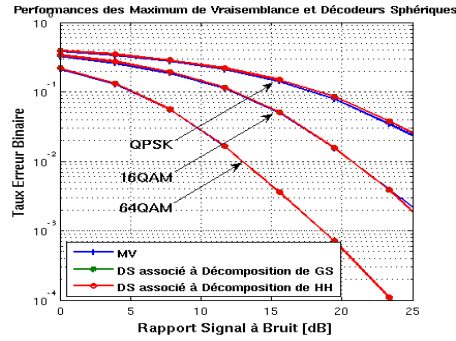


Figure 3. Comparaison des techniques MV, GS associé au DS classique et HH associé au DS adapté. Cas d'un canal MIMO 2×2 de Rayleigh et de modulations QPSK, 16-QAM et 64-QAM respectivement, sans terminaison précoce.

superposées. Les performances en termes de TEB sont très proches, que ce soit pour le MV, le DS classique associé à la décomposition QR de GS ou le DS adapté associé à la décomposition de HH, pour les complexités annoncées.

La méthode proposée permet d'accroître les performances du DS en termes de complexité, tout en étant équivalente au DS associé à un prétraitement de GS et proche du MV en termes de TEB.